

## What Is Claimed Is:

1. A speech processing system for processing an input speech spectrum vector

comprising:

a perceptual speech processor for perceptually processing the input speech spectrum vector to generate a perceptual spectrum;

a storage device for storing a plurality of reference spectrum vectors; and

a phonetic feature mapper, coupled to said perceptual speech processor and to said storage device, for mapping said perceptual spectrum on to said plurality of reference spectrum vectors.

2. The speech processing system of claim 1 wherein said perceptual speech processor comprises:

a masking effector for noise masking the input speech spectrum vector to generate a masked input speech spectrum vector;

a minimum audible field curve renormalizer, coupled to said masking effector, for renormalizing said masked input speech spectrum vector responsive to a minimum audible field to generate a renormalized masked input speech spectrum vector; and

a mel-scale resampler, coupled to said minimum audible field renormalizer, for translating said renormalized masked input speech spectrum vector to mel-scale.

3. The speech processing system of claim 1 wherein said phonetic feature mapper comprises:

a projection similarity generator, coupled to said storage device, for generating a plurality of projection similarity calculations of the input spectrum vector onto said plurality of reference spectrum vectors;

a relative projection similarity generator; coupled to said storage device, for generating a plurality of relative projection similarity calculations of the input spectrum vector onto said plurality of reference spectrum vectors, and

a selector, coupled to said projection similarity generator and to said relative projection similarity generator, for selecting a projection similarity from between said projection similarity generator calculations and said relative projection similarity generator calculations responsive to the relative values of the projection similarity and

relative projection similarity of the input speech spectrum vector on said plurality of reference spectrum vectors.

4. The speech processing system of claim 3 wherein said plurality of reference spectrum vectors is comprised of a plurality of stationary vowels.

5. The speech processing system of claim 4 wherein said plurality of stationary vowels is comprised of nine stationary Mandarin vowels.

6. A speech recognition system for recognizing a sampled speech spectrum vector comprising:

a fast Fourier transform analyzer for generating Fourier transforms of the sampled speech spectrum vector;

a perceptual speech processor, coupled to said fast Fourier transform analyzer, for processing said Fourier transforms to generate a perceptual spectrum;

a storage device for storing a plurality of reference spectrum vectors;

a phonetic feature mapper, coupled to said perceptual speech processor and to said storage device, for mapping said perceptual spectrum to said plurality of reference spectrum vectors, thereby selecting at least one reference vector of greatest similarity to said perceptual spectrum; and

a continuous HMM recognizer, coupled to said phonetic feature mapper, for recognizing said at least one reference vector.

7. The speech recognition system of claim 6 wherein said plurality of reference spectrum vectors is comprised of a plurality of stationary vowels.

8. The speech processing system of claim 7 wherein said plurality of stationary vowels is comprised of nine stationary Mandarin vowels.

9. A method for speech processing an input speech spectrum vector comprising the steps of:

perceptually processing the input speech spectrum vector to generate a perceptual spectrum;

storing a plurality of reference spectrum vectors; and

mapping said perceptual spectrum on to said plurality of reference spectrum vectors.

10. The speech processing method of claim 9 wherein said perceptually processing step further comprises the steps of:

noise masking the input speech spectrum vector to generate a masked input speech spectrum vector;

renormalizing said masked input speech spectrum vector responsive to a minimum audible field to generate a renormalized masked input speech spectrum vector; and

5 translating said renormalized masked input speech spectrum vector to mel-scale.

11. The speech processing method of claim 9 wherein said mapping step further comprises the steps of:

generating a plurality of projection similarity calculations of the input spectrum vector onto said plurality of reference spectrum vectors;

10 generating a plurality of relative projection similarity calculations of the input spectrum vector onto said plurality of reference spectrum vectors, and

selecting a projection similarity from between said projection similarity generator calculations and said relative projection similarity generator calculations responsive to the relative values of the projection similarity and relative projection similarity of the input speech spectrum vector on said plurality of reference spectrum vectors.

12. The speech processing method of claim 11 wherein said plurality of reference spectrum vectors is comprised of a plurality of stationary vowels.

13. The speech processing system of claim 12 wherein said plurality of stationary vowels is comprised of nine stationary Mandarin vowels.

20 14. A method for speech recognition of a sampled input speech spectrum vector, comprising the steps of:

generating Fourier transforms of the sampled input speech spectrum vector utilizing a fast Fourier transform analyzer;

generating a perceptual spectrum by processing said Fourier transforms;

25 storing a plurality of reference spectrum vectors;

mapping said perceptual spectrum onto said plurality of reference spectrum vectors;

selecting at least one reference vector of greatest similarity to said perceptual spectrum; and

30 recognizing said at least one reference vector utilizing a continuous HMM recognizer.

15. The speech recognition system of claim 14 wherein said plurality of reference spectrum vectors is comprised of a plurality of stationary vowels.

16. The speech processing system of claim 15 wherein said plurality of stationary vowels is comprised of nine stationary Mandarin vowels.

Approved for Release by NSA on 09-10-2013 pursuant to E.O. 13526